Learning to Recognize Categories: a Fuzzy Hybrid Approach

Eneida A. Mendonca, MD (*), Altino Pavan, Msc (**), Beatriz dF Leao, MD, PhD (*,**)
(*) Institute of Cardiology RS (ICFUC), Porto Alegre, RS, Brazil
(**) Institute of Informatics, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil

Hybrid systems are being proposed as an efficient solution to build decision support systems. By combining symbolic and connectionist paradigms into a single environment, hybrid system offers the best of two techniques. The ability to acquire knowledge automatically from a case database is supported by the connectionist component, while the explanation of what it has learned and why is supported by the symbolic component. The work described here is a further development in HYCONES - a tightly-coupled Hybrid Connectionist Expert System that integrates the formalism of frames with three different neural networks: the Combinatorial Neural Model (CNM), Semantic ART and Simplified Fuzzy ARTMAP.

The domain chosen to validate this new version was cardiac surgery post-operatory care. To represent domain knowledge HYCONES uses two basic frame's hierarchies: *diagnosis* and *findings*. The findings hierarchy describes the declarative aspects of the domain knowledge and the diagnosis hierarchy stores the possible solutions for the addressed problems. The head of the post-operative Intensive Care Unit of ICFUC was the expert from whom the knowledge was elicited. At the end, 63 diagnoses were described, using a special knowledge acquisition methodology that elicits knowledge as knowledge graphs (KGs). This methodology has a direct translation to HYCONES formalism. Several times during the knowledge acquisition phase, the expert proposed with KGs that used categories to qualify the medical findings, such as low, normal or high arterial blood pressure. To represent this type of data, HYCONES offers to the knowledge engineer the ability to describe quantitative data by a set of linguistic variables. The definition of the membership function of this type of variables came from the expert, through a graphical representation in a X-Y graph. Quantitative data representing, information about hemodynamic parameters, EKG measurements and lab test results were represented as fuzzy variables. At the end of the knowledge acquisition phase, 48 quantitative findings resulted in 122 linguistic variables and their respective membership functions. To represent these fuzzy variables in the symbolic component, the domain knowledge was divided into two sub-classes: quantitative-findings and qualitative-findings. An object to describe the linguistic variable was created, *language-var* which contains the instances of each language variable. A semantic link connects these two objects. When in consultation mode, HYCONES verifies if the patient's finding is a quantitative variable. If so, instead of offering the raw data for processing by the neural networks, the corresponding linguistic variable and value of its membership function (in a 0 to 1 scale) is given to the neural networks.

To validate this approach, a training database was created. To build the database, data from adult patients submitted to cardiac surgery in ICFUC and in Sao Lucas Hospital - Catholic University of RS was collected during a period of 3 months (64 from ICFUC and 34 Sao Lucas Hospital). These data were separated in two groups: 63 patients were randomly assinged to the training set and 35 to the validation set. The assessment of each patient's cardiovascular and respiratory diagnosis was defined as the starting point to this new validation. The predominant diagnoses found in the training set were: marginal cardiac output, adequate cardiac output, low cardiac output, sinus rhythm and pneumonia. After all cases had been collected, the learning-from-scratch algorithm was executed. Through this procedure, a non-operand CNM neural-network was generated. Next, the inductive learning phase was started and, for each neural network arch, the punishment and reward accumulators were calculated. Finally, a normalization and pruning mechanisms eliminated all neural networks connections whose weight was below a pre-set threshold (0.7 in this case).

The system correctly pointed out 65 out 71 diagnoses (91.5%). The semantic contents of the neural nets that pointed out the correct diagnosis were meaningful and very similar to the KGs elicited from the expert.

The results are promising. The system was able to automatically translate quantitative data into a categorical variable figuring out the value of its membership function and deliver this information to the neural networks. It was also able to construct meaningful categories similar to ones pointed out by the expert, with only a few training cases. Presently, the therapeutic knowledge is being inserted into the frames hierarchy and a new validation will be executed.

0195-4210/95/$5.00